Attorney Docket No. <u>YO998-522</u>

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

Applicant(s): L.D. Comerford et al.
Docket No.:   YO998-522
Serial No.:   09/460,913
Filing Date: December 14, 1999
Group: 2654
Examiner: Abul K. Azad

Title: Methods and Apparatus for Contingent
               Transfer and Execution of Spoken
               Language Interfaces

SUPPLEMENTAL APPEAL BRIEF

RECEIVED

MAY 0 4 2004

Technology Center 2600

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This Supplemental Appeal Brief is submitted in response to the Office Action dated January 26, 2004 in the above-referenced application, in which the Examiner reopened prosecution in response to the Appeal Brief filed October 21, 2003. The arguments set forth in the Appeal Brief are incorporated by reference herein.

Applicants (hereinafter "Appellants") have submitted concurrently herewith a response to the Office Action, requesting reinstatement of the appeal pursuant to 37 C.F.R. §1.193(b)(2).

Appellants note that in the present Office Action, the Examiner has raised, for the first time, objections to the specification. Specifically, the Examiner has objected to the specification contending that it contains references to related applications which should be made by application number and filing date. The Examiner has also objected to the specification under MPEP §608.01 as containing an embedded hyperlink and/or other form of browser-executable code. Appellants

request that these objections be held in abeyance pending a resolution of the appeal, at which point Appellants will amend the specification as may be required to address the objections.

REAL PARTY IN INTEREST

The present application is assigned to International Business Machines Corporation (IBM Corp.), as evidenced by an assignment recorded April 3, 2000 in the U.S. Patent and Trademark Office at Reel 010671, Frame 0264. The assignee IBM Corp. is the real party in interest.

RELATED APPEALS AND INTERFERENCES

There are no known related appeals and interferences.

STATUS OF CLAIMS

Claims 1-19 are pending in the present application. Claims 1-5, 9 and 10 stand rejected under 35 U.S.C. §102(b). Claims 6, 7, 13 and 19 stand rejected under 35 U.S.C. §102(e). Claims 8, 11 and 14-18 stand rejected under 35 U.S.C. §103(a). Claims 1-19 are appealed.

STATUS OF AMENDMENTS

There have been no amendments filed subsequent to the final rejection.

SUMMARY OF INVENTION

The present invention relates to "methods and apparatus for contingent transfer and execution of spoken language interfaces" (Specification; page 2, lines 24-25). A spoken language interface is defined in sets of user interface files. The user interface files are referred to as "vocabularies files, prompt files, profiles and scripts depending on the role they play in structuring the interface" (Specification; page 3, lines 4-5). As used by the present invention, the term "spoken language interface" is intended to refer to the general act of speaking to a machine, listening to a machine, and/or interacting with a machine through utterances or audible expressions, and does **not** refer to a particular lingual type (e.g., English or Spanish). The present invention provides a means by which

a spoken language interface, such as a user interface data set and a dialog manager engine capabilities of a personal speech assistant (PSA), can be managed based on certain contingencies (Specification; page 3, lines 24-26).

An important aspect of the invention is its ability to <u>dynamically instantiate a new application and its corresponding spoken language interface</u> (Specification; page 42, lines 5-7). It should be appreciated that the spoken language user interface is a <u>collection of operable features</u> that allows a user to interact with the application. For example, user utterances may operate the features of the application, e.g., by supplying a reference to one or more events to be processed by the target application (Specification; page 3, lines 5-6). The term "event" is used by the present invention in a conventional sense in the context of event handling programs. Event handling is a feature of the application. These operable features, which are built into an application and are controlled at least in part by user utterances, are to be <u>distinguished from data</u> on which the application program acts.

In an illustrative embodiment of the invention, as shown in FIG. 17 of the drawings, a user is assumed to have added a new speech aware application to a personal digital assistant (PDA), which is in communication with a PSA, and chooses to access that application (17000) (Specification; page 39, lines 1-2). As the application begins execution for the first time, it accesses a Voice Library (17010) provided to the PDA which recognizes the application as a new application and sends a notification to the PSA (17020) that the new application has been received (Specification; page 39, lines 4-8). The PSA, receiving the message (17030), determines that it has no user interface data structure entries corresponding to the new application (Specification; page 39, lines 9-11). The PSA, finding that the user interface files are not in place (17040), send a request to the PDA for the appropriate user interface data set corresponding to the new application (Specification; page 39, lines 14-15). On receiving the user interface data (17070), the PSA dialog manager, continuing the "hard-coded" function of adding a new application, places the name of the new application in the application list (17080), sets the user interface files of the new application as the active application files (17090), and sends a message indicating that the user may now speak to the new application (17100) (Specification; page 39, lines 20-27).

3

Application Serial No. 09/460,913

Independent claim 1 specifies, in an apparatus for providing a portable spoken language interface for a user to a device in communication with the apparatus, a method for modifying a data structure containing at least one user interface data set. The method includes adding a new application to the device, generating a second user interface data set corresponding to the new application, transferring the second user interface data set from the device to the apparatus, and loading the second user interface data set into the data structure of the apparatus. The second user interface data set represents spoken language interface elements and data recognizable by the new application. Independent claims 6, 16 and 19 specify a method, apparatus and article of manufacture, respectively, for automatically providing contingent transfer and execution of one or more spoken language interfaces for a user with respect to at least one external network with which the user interacts.

ISSUES PRESENTED FOR REVIEW

1. Whether claims 1-5, 9 and 10 are properly rejected under 35 U.S.C. §102(b) as being anticipated by U.S. Patent No. 5,752,232 to Basore et al. (hereinafter "Basore").

2. Whether claims 6, 7, 13 and 19 are properly rejected under 35 U.S.C. §102(e) as being anticipated by U.S. Patent Application Publication No. US 2002/0091513 A1 to Mozer et al. (hereinafter "Mozer").

3. Whether claims 11, 12, 14 and 15 are properly rejected under 35 U.S.C. §103(a) as being unpatentable over Basore and Mozer, and further in view of U.S. Patent No. 6,144,938 to Surace et al. (hereinafter "Surace").

4. Whether claims 8, 16 and 17 are properly rejected under 35 U.S.C. §103(a) as being unpatentable over Mozer in view of U.S. Patent No. 5,481,616 to Freadman (hereinafter "Freadman").

5. Whether claim 18 is properly rejected under 35 U.S.C. §103(a) as being unpatentable over Mozer and Freadman, in further view of U.S. Patent No. 6,044,347 to Abella et al. (hereinafter "Abella").

GROUPING OF CLAIMS

Issues 1-4 above each involve a ground of rejection which Appellants contest and which applies to a group of two or more claims.

With regard to Issue 1, claims 1-5, 9 and 10 do not stand or fall together. Claims 1, 2 and 9 stand or fall together, claims 3-5 stand or fall together and are believed to be separately patentable, and claim 10 is believed to be separately patentable.

With regard to Issue 2, claims 6, 7, 13 and 19 do not stand or fall together. Claims 6, 7 and 19 stand or fall together, and claim 13 is believed to be separately patentable.

With regard to Issue 3, claims 11, 12, 14 and 15 do not stand or fall together. Claims 11 and 14 stand or fall together, and claims 12 and 15 stand or fall together and are believed to be separately patentable.

With regard to Issue 4, claims 8, 16 and 17 stand or fall together.

ARGUMENT

Issue 1

A.   Claims 1, 2 and 9

As stated above, independent claim 1 specifies, in an apparatus for providing a portable spoken language interface for a user to a device in communication with the apparatus, a method for modifying a data structure containing at least one user interface data set. The method includes adding a new application to the device, generating a second user interface data set corresponding to the new application, transferring the second user interface data set from the device to the apparatus, and loading the second user interface data set into the data structure of the apparatus. The second user interface data set represents spoken language interface elements and data recognizable by the new application.

With regard to claim 1, the Examiner contends that Basore discloses, in an apparatus for providing a portable spoken language interface for a user to a device in communication with the apparatus, a method including all of the limitations set forth in the claim 1 (Office Action dated January 26, 2004; page 3, paragraph 4 to page 7, paragraph 3). Basore is directed to "a voice

activated device using speaker independent speech recognition which is capable of being set to establish a connection automatically at a pre-specified time to a remote location. Upon establishing the connection, phonetic spellings needed for speech recognition, as well as additional application data, are communicated in digitized form to the voice activated device from the remote location and stored in the device" (Basore; column 1, lines 49-52). For at least the reasons set forth below, Appellants respectfully disagree with the Examiner's contention that claim 1 is anticipated by Basore.

Applicants submit that claim 1 is patentable over the Basore reference. Specifically, Basore fails to teach or suggest at least a dialog manager, as recited in claim 1. In the Office Action, the Examiner incorrectly analogizes the application software running in the microprocessor and DSP chip disclosed in Basore to the dialog manager of the claimed invention (Office Action dated January 26, 2004; page 4, paragraph 2). While Appellants may acknowledge that the terminology used in claim 1 need not be identical to the terminology disclosed in the prior art, in order to sustain a rejection under 35 U.S.C. §102, the prior art reference must at least "teach every aspect of the claimed invention either explicitly or impliedly. Any feature not directly taught must be inherently present" (MPEP §706.02; emphasis added). Appellants submit that Basore fails to teach or suggest at least a dialog manager having the functionality set forth in the subject claims. Furthermore, the claimed features are not inherently present in Basore.

The dialog manager recited in claim 1 is clearly defined in the specification. It is well-settled that "[a] patentee is his own lexicographer"(Canaan Prod., Inc. v. Edward Don & Co., 388 F.2d 540, 544, 156 USPQ 295, 298 (7th Cir. 1968)), and that "[a] patentee may define his own terms, regardless of common or technical meaning, and fairness to the patentee requires the court to accept his definition of words, phrases, and terms" (International Cork Co. v. New Process Cork Co., 6 F.2d 420, 422 (2d Cir. 1925)(emphasis added)). The Federal Circuit has reiterated this doctrine, stating that: "It is a well-established axiom in patent law that a patentee is free to be his or her own lexicographer, . . . and thus may use terms in a manner contrary to or inconsistent with one or more of their ordinary meanings. For this reason, an analysis of the specification and prosecution history

is important to proper claim construction." *Hormone Research Foundation, Inc. v. Genetech, Inc.*, 904 F.2d 1558, 1563, 15 USPQ2d 1039, 1043 (Fed. Cir. 1990) (emphasis added).

In contrast to the claimed invention, the application software in Basore has no <u>dialog management</u> functionality, as the term is defined in the present specification, and thus cannot reasonably be analogized to a dialog manager. The dialog manager recited in claim 1 is explicitly defined in the present specification to include "an interpreter component 1010, an engine manager 1020 and a serial port communications interface 1030" (present specification; page 7, lines 19-21). When the specification provides definitions for terms appearing in the claims, the specification can be used in interpreting such claim language (MPEP §2111.01; citing *In re Vogel*, 422 F.2d 438, 441, 164 USPQ 619, 622 (CCPA 1970)). Basore fails to teach or suggest that the software running on the microprocessor or DSP implements such dialog manager functionality. In this regard, Appellants assert that the Examiner's characterization of Basore inappropriately attempts to impute a much broader scope to the Basore reference than what is specifically taught or suggested by the reference itself.

The apparatus taught by Basore is a dedicated appliance running a <u>single application</u> for retrieving and presenting predetermined information to a user in response to recognized voice commands uttered by the user. Basore makes no reference at all to the concept of dialog management in general (i.e., communication between two or more applications coexisting on the same platform). Additionally, since the voice activated device disclosed in Basore performs <u>no resource allocation</u>, there is no reason to include a dialog manager, or any dialog management functionality, for handling multiple applications.

In contrast to Basore, the engine manager, a component of the dialog manager recited in claim 1, <u>provides resource allocation between engines and/or processes</u> (i.e., applications). For example, the present specification, on page 14, lines 1-3 states: "The engine manager 3020 determines whether or not any resources needed by the command recognition engine are in use by other engines or processes. If they are, it halts those activities." This resource allocation capability, which is a function of the dialog manager, is not taught or suggested by the prior art of record. As

previously stated, Basore does not require resource allocation since the voice activated device disclosed in Basore runs a <u>single dedicated application program</u>.

In addition to failing to teach or suggest a dialog manager, Basore also fails to teach or suggest "at least one <u>user interface data set</u> . . . representing spoken language interface elements and data <u>recognizable by the application of the device</u>," as set forth in claim 1 (emphasis added). In this regard, the Examiner incorrectly analogizes "a phonetic acoustic models dictionary" and "stored application data phonetic spelling used in a particular application" disclosed in Basore with the user interface data set of the claimed invention (Office Action dated January 26, 2004; page 4, last paragraph). Phonetic spellings cannot reasonably be analogized to spoken language interface elements, as required by the claimed invention, since they <u>do not provide a user interface functionality</u> with an application associated with an external device. Such phonetic spellings taught by Basore merely define a list of recognized words.

Basore states that "[t]he phonetic acoustic models database 126 stores a plurality of <u>models of how phonemes are spoken</u>" (Basore; column 2, lines 64-65). Moreover, Basore states that "it is assumed that <u>the additional application data comprises the responses</u> which are also stored in the memory unit 125" (Basore; column 4, lines 46-48; emphasis added). However, neither the "phonetic spellings" nor the "additional application data" taught by Basore provide any command recognition or user interface functionality to the device, among other things, and thus cannot be employed as a user interface data set as recited in claim 1.

Unlike Basore, the user interface data set recited in claim 1 provides "a spoken language interface to <u>applications</u>" (Specification; page 17, line 28, to page 18, line 1; emphasis added). Here, the term "applications," as is conventionally understood, refers to one or more software programs or processes residing on the device, each program or process executing distinct code. Basore uses a nonanalogous definition of the term "application." Consequently, the prior art of record fails to teach or suggest "a user interface data set," as required by claim 1.

Because the voice activated device disclosed in Basore fails to disclose at least a dialog manager or dialog management functions, as stated above, it follows that Basore also fails to teach or suggest many of the other corresponding functionalities of the spoken language interface

apparatus recited in claim 1. For example, Basore fails to disclose that "the dialog manager provides the spoken language interface element associated data to the application of the device for processing in accordance therewith," and that "the application of the device, on processing that element, provides a reference to an interface element to be spoken," as required by claim 1.

Basore further fails to disclose the steps of adding a new application to the device and generating a second user interface data set corresponding to the new application, as required by claim 1. With regard to the former step, the Examiner contends that such a step is disclosed in Basore at column 4, lines 22-29 (Office Action dated January 26, 2004; page 6, paragraph 5). Specifically, Basore states: "Applications for which phonetic spellings and additional application data may be communicated to the voice-activated device 120 include, for example, information related to weekly television schedules, daily weather reports, . . ." (Basore; column 4, lines 22-29). Appellants submit, however, that the "applications" to which Basore refers are merely the specific uses to which the voice activated device may be put, and are not new software programs running on the device. The "phonetic spellings" in Basore are the subset of recognized words used in the particular application (Basore; column 3, lines 40-42). For example, Basore, with reference to FIG. 3, illustrates how the voice activated device can be used in a television guide application for remotely obtaining program listing information (Basore; column 4, lines 38-40).

While Basore may disclose remotely downloading a new vocabulary including phonetic spellings for recognizing a new subset of words relating to a specific application for which the device is to be used, the voice activated device itself is still capable of performing only a dedicated function. Basore fails to teach or suggest any means for adding another application to the device, as required by the claimed invention. Furthermore, Basore fails to teach or suggest generating a second user interface data set, as required by claim 1. In this regard, the Examiner contends (Office Action dated January 26, 2004; page 6, last paragraph) that such a step is disclose in Basore at column 4, lines 1-21 as "additional application data that is relevant to the applications for whose services the user is registered" (Basore; column 4, lines 7-10; emphasis added). These additional applications, however, are not applications running on the device, but rather are remote applications which the voice activated device of Basore may access for obtaining requested information, e.g., television program

9

listings, daily weather reports, stock prices, etc. (Basore; column 4, lines 24-27). Basore is thus clearly distinguishable from the claimed invention, which requires adding a new application to the device itself.

For at least the foregoing reasons, Appellants assert that claim 1 is patentable over the cited prior art.

Claims 2 and 9 depend from claim 1 and are therefore believed to be patentable for at least the reasons set forth above with respect to claim 1.

B.   Claims 3-5

Claim 3 further specifies "removing a user interface data set from the data structure." With regard to claim 3, the Examiner contends that such feature is disclosed in Basore at column 4, lines 15-21 as "additional application data that have been added, deleted or changed since the previous update are communicated to the device" (Office Action dated January 26, 2004; page 7, paragraph 5). Appellants respectfully disagree with this contention.

As noted above with regard to claim 1, the user interface data set recited in the subject claims is not analogous to the "application data" set forth in Basore. While the additional application data taught by Basore may be "relevant to the particular application" (Basore; column 3, lines 42-44), it does not provide user interface functionality to the device, and is thus not analogous to the user interface data set recited in the subject claims. Morover, Basore fails to disclose adding or deleting user interface data sets. Rather, in contrast to the claimed invention, Basore discloses deleting obsolete phonetic spellings and additional application data in order to preserve memory (Basore; column 4, lines 15-21), which is not analogous to removing a user interface data set, as required by claim 3.

For at least the reasons set forth above, Appellants assert that claim 3 is separately patentable over the cited prior art.

Claims and 4 and 5 depend from claim 3 and are therefore believed to be patentable for at least the reasons set forth above with respect to claim 3.

C.   Claim 10

Claim 10 further defines the method for modifying a data structure containing at least one user interface data set as prompting the user for information comprising a spoken utterance and the device manager "modifying at least one of a predetermined parameter of the device and an application running on the device" in response to the spoken utterance. The prior art of record fails to teach or suggest at least this feature of the claimed invention.

As per claim 10, the Examiner contends that this additional feature is disclosed in Basore, at column 4, lines 22-36 as "updating phonetic spelling and additional application data" (Office Action dated January 26, 2004; page 8, paragraph 3). Appellants respectfully disagree with this contention and submit that while Basore may disclose modifying the list of words to be recognized (phonetic spellings) or modifying the responses to those words (additional application data), Basore fails to teach or suggest modifying one or more operating parameters of the device and/or modifying the application itself in response to the spoken utterance, as explicitly set forth in claim 10. Rather, the device taught by Basore functions in essentially the same dedicated manner, regardless of the words that are contained in its dictionary (127) or the responses that are stored in its memory unit (125) (Basore; FIG. 1).

Inasmuch as the prior art of record clearly fails to teach or suggest the above-noted limitations, claim 10 is believed to be independently patentable.

Issue 2

Appellants hereby re-allege and incorporate by reference the arguments relating to Issue 1 above in their entirety.

A.   Claims 6, 7 and 19

As was noted above, independent claims 6 and 19 specify a method and article of manufacture, respectively, for automatically providing contingent transfer and execution of one or more spoken language interfaces for a user with respect to at least one external network with which the user interacts. The method and article of manufacture further include the steps of requesting a

11

spoken language interface data set from the external network upon discovery of the external network, transferring the spoken language interface data set to portable spoken language interface device associated with the user, and loading the spoken language interface data set into a data structure of the device for use by the user interfacing with the external network.

With regard to claims 6 and 19, the Examiner contends that Mozer teaches "a method of automatically providing a spoken language interface for a user with respect to at least one external network with which the user interacts, wherein the user process [sic] a portable spoken language interface device having a data structure for storing one or more user interface data sets used to provide one or more spoken language interfaces" including each of the steps recited in claims 6 and 19 (Office Action dated January 26, 2004; page 9, first paragraph). Appellants respectfully disagree with this contention.

Appellants submit that claims 6 and 19 are patentable over the Mozer reference. Specifically, Mozer fails to disclose a method or article of manufacture, respectively, capable of automatically providing a spoken language interface for a user with respect to at least one external network with which the user interacts, as required by the subject claims. In this regard, Appellants assert that the recognition set and weight set taught by Mozer, on page 3, paragraph 0032, are not analogous to the spoken language interface data set recited in the claimed invention and defined by the present specification. The recognition set disclosed in Mozer is merely a set of words recognized by the dedicated interface associated with the speech recognition system. While Mozer may disclose that the recognition system can select a new set of words and associated weights for recognizing a new user utterance (Mozer; page 3, paragraph 0033), Mozer fails to teach or suggest automatically providing a new spoken language interface for the user, as set forth in claims 6 and 19. Moreover, Mozer fails to teach or suggest dynamically changing an application of the speech recognition system. Rather, Mozer discloses a system that is dedicated to a single application (e.g., a compact disc changer application).

The present specification provides a clear definition of a "spoken language interface data set" as set forth in the subject claims. The specification, at least at page 3, beginning at line 3, explicitly defines a spoken language interface data set as comprising:

12

... sets of user interface files. These are referred to as vocabularies files, prompt files, profiles and scripts depending on the role they play in structuring the interface. Vocabulary files provide tables relating possible user utterances with events and data that are understood by the target application. Prompt files provide sets of standard responses that may be used by applications to prompt the user to actions or inform the user about application status. Profiles set the parameters for the operation of the PSA hardware such as the voicing parameters used by the text to speech engine. . . Scripts provide instruction programming sequences of PSA services.

It is well-settled that "[a] patentee is his own lexicographer." *Canaan Prod., Inc. v. Edward Don & Co.*, 388 F.2d 540, 544, 156 USPQ 295, 298 (7th Cir. 1968). The Federal Circuit has reiterated this doctrine, stating that: "It is a well-established axiom in patent law that a patentee is free to be his or her own lexicographer, . . . and thus may use terms in a manner contrary to or inconsistent with one or more of their ordinary meanings. For this reason, an analysis of the specification and prosecution history is important to proper claim construction." *Hormone Research Foundation, Inc. v. Genetech, Inc.*, 904 F.2d 1558, 1563, 15 USPQ2d 1039, 1043 (Fed. Cir. 1990).

The recognition and weight sets disclosed in Mozer, in contrast to the claimed invention, are not used to structure a spoken language interface itself, nor are they able to define parameters of operation of the speech recognition system. Therefore, such recognition and weight sets are clearly not analogous to the spoken language interface data set recited in the claimed invention. To further highlight the differences between the two different types of data sets, from a system hierarchy standpoint, the recognition set and associated weight set taught by Mozer reside in a device/driver layer of the system, while the spoken language interface data set of the claimed invention resides in the operating system shell layer, which is a higher hierarchical layer of the system.

For at least the foregoing reasons, Appellants assert that claims 6 and 19 are patentable over the cited prior art.

Claim 7 depends from claim 6 and is therefore believed to be patentable for at least the reasons set forth above with respect to claim 6.

B.   Claim 13

With regard to claim 13, this claim further defines the method of automatically providing a spoken language interface as prompting the user for information comprising a spoken utterance and "modifying at least one of a predetermined parameter of the device and an application running on the device" in response to the spoken utterance. The prior art of record fails to teach or suggest at least this feature of the claimed invention.

As per claim 13, the Examiner contends that such feature is taught by Mozer on page 3, paragraph 0029, where it states that "[t]he information presented to the user may include prompts for input to microphone 105 or application specific information" (Office Action dated January 26, 2004; page 10, paragraph 2). Appellants assert that merely providing prompts for a specific application is well-known in the art. However, this is not an accurate characterization of the invention set forth in claim 13. Specifically, Mozer fails to teach or suggest that the system is capable of modifying one or more parameters of the system and/or an application running on the system in response to the user utterance, as expressly required by claim 13. While Mozer may disclose that "pattern recognition system 112 receives the recognition result and selects a new set of words and associated weight set based on this result" (Mozer; page 3, paragraph 0033), Appellants assert that merely selecting a new set of words to be recognized does not amount to modifying one or more operating parameters of the system. The pattern recognition system and related components of Mozer continue to operate in the same predefined manner.

Inasmuch as Mozer clearly fails to teach or suggest the above-noted limitations, claim 13 is believed to be independently patentable.

Issue 3

Appellants hereby re-allege and incorporate by reference the arguments relating to Issues 1 and 2 above in their entirety.

A.   Claims 11 and 14

Claims 11 and 14 depend from claims 1 and 6, respectively, and are therefore believed to be patentable for at least the reasons set forth above with respect to claims 1 and 6. Furthermore, these claims are independently patentable for at least the reasons set forth below.

Claim 11, like claim 14, further defines the step of prompting the user for information as "storing one or more user experience parameters corresponding to a familiarity of the user with a predetermined procedure of the application; and selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on experience parameters, the selected prompt substantially matching the stored experience parameters of the user." Appellants submit that the prior art of record clearly fails to teach or suggest at least this feature of the claimed invention.

As per claims 11 and 14, the Examiner acknowledges that Basore and Mozer fail to teach the above-noted additional features (Office Action dated January 26, 2004; page 11, first paragraph). However, the Examiner contends that such features are disclosed in Surace (Office Action dated January 26, 2004; page 11, paragraph 3). Appellants respectfully disagree with this contention.

Appellants respectfully assert that the recited combinations of Basore, Mozer and Surace fail to establish a *prima facie* case of obviousness under 35 U.S.C. §103(a), as specified in M.P.E.P. §2143. As set forth therein, M.P.E.P. §2143 states that in order to "establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation . . . to modify the reference or combine reference teachings. Second, there must be some reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations." All three of these basic criteria must be met in order to sustain a *prima facie* case of obviousness.

Appellants assert that there lacks the requisite motivation or suggestion, either in the references themselves or in the knowledge generally available to one skilled in the art, to combine or modify the teachings of Basore or Mozer in view of Surace to arrive at the claimed invention. Basore discloses a voice activated device for providing access to remotely retrieved data (Basore; column 1, lines 49-52), and Mozer teaches an inexpensive speech recognition system for providing

speech recognition capabilities to consumer electronic products (Mozer; page 2, paragraphs [0011]-[0012]). Surace, on the other hand, is directed generally to user interfaces, and more particularly, to a voice user interface which "behaves consistently with social and emotional norms" (Surace; column 1, lines 50-51).

Appellants assert that Basore, Mozer and Surace are directed to different fields of art, seek solutions to entirely different problems, and, as such, "particular findings must be made as to the reason the skilled artisan, with no knowledge of the claimed invention, would have selected these components for combination" (In re Kotzab, 217 F.3d 1365, 1371, 55 USPQ2d 1313, 1317 (Fed. Cir. 2000); emphasis added).

Although not necessarily dispositive of nonanalogous art, a lack of motivation or suggestion to combine the references is further evidenced by the fact that Basore, Mozer and Surace all have dissimilar patent classifications. Moreover, the mere fact that the references can be combined or modified does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination. In re Mills, 916 F.2d 680, 16 USPQ2d 1430 (Fed. Cir. 1990) Appellants assert that no such suggestion exists in either the Basore or Mozer references or the Surace reference. Thus, Appellants submit that one skilled in the art would not look to a user interface methodology (Surace) to find inspiration for improving or otherwise modifying a voice activated device for obtaining remotely retrieved data (Basore) or a speech recognition system for adding speech recognition functionality to consumer electronic products (Mozer), or vice versa.

For at least the above reasons, a prima facie case of obviousness has not been established. "Obviousness cannot be established by combining the teachings of the prior art to produce the claimed invention, absent some teaching, suggestion or incentive supporting the combination"(In Re Bond, 910 F.2d 831, 833, 15 USPQ2d 1566, 1568 (Fed. Cir. 1990)). The prior art of record provides no such teaching, suggestion, or incentive supporting the combination of Basore and Mozer with Surace.

For at least the above reasons, claims 11 and 14 are believed to be separately patentable.

B.   Claims 12 and 15

Claims 12 and 15 depend from claims 1 and 6, respectively, and are therefore believed to be patentable for at least the reasons set forth above with respect to claims 1 and 6. Furthermore, as stated above in connection with claims 11 and 14, Appellants believe that there lacks the relevant motivation or suggestion to combine the Basore, Mozer and Surace references. Consequently, claims 12 and 15 are believed to be separately patentable.

Issue 4

Appellants hereby re-allege and incorporate by reference the arguments relating to Issues 1-3 above in their entirety.

As was noted above, independent claim 16 specifies an apparatus for automatically providing contingent transfer and execution of one or more spoken language interfaces for a user with respect to at least one external network with which the user interacts. The apparatus includes "a portable spoken language interface device; and a personal data assistant (PDA) operatively coupled to the spoken language interface device, the PDA including at least one application associated therewith." The portable spoken language interface device is operative to "(i) request a spoken language interface data set from the external network upon discovery of the external network; (ii) receive from the external network the spoken language interface data set; and (iii) load the spoken language interface data set into the data structure of the portable spoken language interface device for use by the user interfacing with the external network."

With regard to claim 16, the Examiner acknowledges that Mozer fails to teach a personal data assistant operatively coupled to a spoken language interface device (Office Action dated January 26, 2004; page 12, last paragraph). However, the Examiner contends that Freadman discloses an analogous apparatus including a microphone, speech input and output, and a speech interface to another device, coupled to a personal digital assistant (Office Action dated January 26, 2004; page 13, first paragraph). Appellants respectfully disagree with the Examiner's contention that the combination of Mozer and Freadman discloses all of the limitations recited in claim 16.

First, Appellants assert that there lacks the requisite motivation or suggestion, either in the prior art of record or in the knowledge generally available to one skilled in the art, to modify the teachings of Mozer in view of Freadman in order to sustain a *prima facie* obviousness rejection of claim 16. Specifically, Mozer is directed to an inexpensive <u>speech recognition system</u> for providing speech recognition capabilities to consumer electronic products (Mozer; page 2, paragraphs [0011]-[0012]), while Freadman, by contrast, is directed to a <u>plug-in sound accessory</u> for equipping a portable computer with the capability to produce sound (Freadman; column 1, lines 10-11; emphasis added). Freadman states that "this invention proposes . . . adding additional circuitry (see FIG. 3) to a conventional flash card so as to obtain an enhanced PCMCIA card 22 operative, together with a master speaker module 24, to equip the computer with a sound capability" (Freadman; column 3, lines 6-11). Mozer and Freadman are directed to entirely different fields of art, seek solutions to entirely different problems, and, as such, "particular findings must be made as to the reason the skilled artisan, <u>with no knowledge of the claimed invention</u>, would have selected these components for combination" (*In re Kotzab*, 217 F.3d 1365, 1371, 55 USPQ2d 1313, 1317 (Fed. Cir. 2000); emphasis added).

Appellants submit that Mozer and Freadman are nonanalogous art and are therefore not combinable for the purpose of sustaining a *prima facie* case of obviousness against the subject claims. The mere fact that the portable spoken language interface device <u>may</u> be combined with a PDA does not, by itself, render the claimed combination obvious. It is well-settled law that "[o]bviousness cannot be established by combining the teachings of the prior art to produce the claimed invention, absent some teaching or suggestion supporting the combination. Under section 103, teachings of references can be combined *only* if there is some suggestion or incentive to do so." *ACS Hosp. Systems, Inc. v. Montefiore Hosp.*, 732 F.2d 1572, 1577, 221 USPQ 929, 933 (Fed. Cir. 1984). Appellants assert that one skilled in the art would not look to a plug-in sound accessory for a portable computer, as taught by Freadman, to find inspiration for improving or otherwise modifying a speech recognition system for adding speech recognition functionality to consumer electronic products, as taught by Mozer, or vice versa.

For at least the above reasons, Appellants assert that a *prima facie* case of obviousness has not been established.

Second, Appellants submit that, even assuming, *arguendo*, that the Mozer and Freadman references can be combined, the combination of Mozer and Freadman fails to disclose all of the limitations set forth in claim 16. For example, Mozer fails to teach or suggest a portable spoken language interface device that is operative to request a spoken language interface data set from an external network upon discovery of the external network, as required by claim 16. Additionally, since Mozer relates to providing inexpensive speech recognition for a <u>dedicated appliance</u> that does not otherwise have such capability (Mozer; Abstract), there is no reason as to why Mozer would require requesting the spoken language interface data set from the external network, receiving from the external network the spoken language interface data set, and loading the new spoken language interface data set into the data structure of a portable spoken language interface device for use by the user interfacing with the external network, as required by claim 16. Freadman fails to supplement the deficiencies of Mozer in this regard.

For at least the foregoing reasons, Appellants assert that claim 16 is patentable over the cited prior art.

Claim 8 depends from claim 6 and is therefore believed to be patentable for at least the reasons set forth in Issue 2 above with respect to claim 6.

Claim 17 depends from claim 16 and is therefore believed to be patentable for at least the reasons set forth above with respect to claim 16.

Issue 5

Appellants hereby re-allege and incorporate by reference the arguments relating to Issues 1-4 above in their entirety.

Claim 18 further specifies that "the portable spoken language interface device program comprises a personal speech assistant (PSA)" including, among other things, a dialog manager and at least one user interface data set operatively coupled to the dialog manager. Appellants respectfully disagree with the Examiner's contention that Abella supplements the deficiencies of Mozer and

Application Serial No. 09/460,913

Freadman. Abella, for example, fails to teach or suggest a dialog manager having the functionality as defined by the present specification. Furthermore, Appellants assert that claim 18 depends from claim 16 and is therefore believed to be patentable for at least the reasons set forth above with respect to claim 16.

Respectfully submitted,

Wayne L. Ellenbogen
Attorney for Applicant(s)
Reg. No. 43,602
Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560
(516) 759-7662

Date: April 26, 2004

APPENDIX

1. In apparatus for providing a portable spoken language interface for a user to a device in communication with the apparatus, the device having at least one application associated therewith, the spoken language interface apparatus comprising: (A) an audio input system for receiving speech data provided by the user; (B) an audio output system for outputting speech data to the user; (C) a speech recognition engine for generating an output in response to spoken utterances; (D) a speech synthesizing engine for generating a synthesized speech output in response to text data; (E) a dialog manager operatively coupled to the device, the audio input system, the audio output system, the speech recognition engine and the speech synthesizing engine; and (F) at least one user interface data set operatively coupled to the dialog manager, the user interface data set representing spoken language interface elements and data recognizable by the application of the device; wherein: (i) the dialog manager enables connection between the input audio system and the speech recognition engine such that the spoken utterance provided by the user is provided from the input audio system to the speech recognition engine; (ii) the output generated by the speech recognition engine is returned to the dialog manager; (iii) the dialog manager uses the output generated by the speech recognition engine to search the user interface data set for a corresponding spoken language interface element and data which is returned to the dialog manager when found; (iv) the dialog manager provides the spoken language interface element associated data to the application of the device for processing in accordance therewith; (v) the application of the device, on processing that element, provides a reference to an interface element to be spoken; (vi) the dialog manager enables connection between the audio output system and the speech synthesizing engine such that the speech synthesizing engine which, accepting data from that element, generates a synthesized output that expresses that element; and (vii) the audio output system audibly presenting the synthesized output to the user; a method for modifying a data structure containing the at least one user interface data set, comprising:

adding a new application to the device;

21

generating a second user interface data set in accordance with the new application, the second user interface data set representing spoken language interface elements and data recognizable by the new application;

transferring the second user interface data set from the device to the apparatus; and

loading the second user interface data set into the data structure of the apparatus.

2. The method of claim 1, further comprising the step of audibly notifying the user that the new application is useable via the audio output system.

3. The method of claim 1, further comprising the step of removing a user interface data set from the data structure.

4. The method of claim 3, wherein the user interface data set is removed prior to the loading step in accordance with a least recently used algorithm.

5. The method of claim 3, wherein the user interface data set is removed in accordance with a request by an application.

6. A method of automatically providing a spoken language interface for a user with respect to at least one external network with which the user interacts, wherein the user possesses a portable spoken language interface device having a data structure for storing one or more user interface data sets used to provide one or more spoken language interfaces, the method comprising the steps of:

the device requesting a spoken language interface data set from the external network upon discovery of the external network;

the external network transferring the spoken language interface data set to the device; and

loading the spoken language interface data set into the data structure of the device for use by the user interfacing with the external network.

7. The method of claim 6, wherein the device is in wireless communications with the external network.

8. The method of claim 6, wherein the device comprises a personal digital assistant.

9. The method of claim 1, wherein the new application comprises a speech aware application, the speech aware application being responsive to user utterances for at least partially interacting with the new application.

10. The method of claim 1, further comprising the step of:

the device prompting the user for information comprising a spoken utterance, the device manager being responsive to the spoken utterance for operatively modifying at least one of a predetermined parameter of the device and an application running on the device.

11. The method of claim 10, wherein the step of prompting the user for information includes the steps of:

storing one or more user experience parameters corresponding to a familiarity of the user with a predetermined procedure of the application; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on experience parameters, the selected prompt substantially matching the stored experience parameters of the user.

12. The method of claim 10, wherein the step of prompting the user for information includes the steps of:

storing an internal data set including at least one of a date, a time and a number of times which a predetermined procedure of an application is performed; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on information included in

23

the internal data set, the selected prompt substantially matching the stored internal data set.

13. The method of claim 6, further comprising the step of:

the portable spoken language interface device prompting the user for information comprising a spoken utterance, the device being responsive to the spoken utterance for operatively modifying at least one of a predetermined parameter of the device and an application running on the device.

14. The method of claim 13, wherein the step of prompting the user for information includes the steps of:

storing one or more user experience parameters corresponding to a familiarity of the user with a predetermined procedure of the application; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on experience parameters, the selected prompt substantially matching the stored experience parameters of the user.

15. The method of claim 13, wherein the step of prompting the user for information includes the steps of:

storing an internal data set including at least one of a date, a time and a number of times which a predetermined procedure of an application is performed; and

selecting a prompt from a set of prompts for presentation to the user, the set of prompts including varying amounts of instruction based at least in part on information included in the internal data set, the selected prompt substantially matching the stored internal data set.

16. Apparatus for automatically providing contingent transfer and execution of one or more spoken language interfaces for a user with respect to at least one external network with which the user interacts, the apparatus comprising:

a portable spoken language interface device; and

24

a personal data assistant (PDA) operatively coupled to the spoken language interface device, the PDA including at least one application associated therewith;

wherein the portable spoken language interface device is operative to: (i) request a spoken language interface data set from the external network upon discovery of the external network; (ii) receive from the external network the spoken language interface data set; and (iii) load the spoken language interface data set into the data structure of the portable spoken language interface device for use by the user interfacing with the external network.

17. The apparatus of claim 16, wherein the portable spoken language interface device is in wireless communication with the external network.

18. The apparatus of claim 16, wherein the portable spoken language interface device comprises a personal speech assistant (PSA), the PSA comprising:

an audio input system for receiving speech data provided by the user;

an audio output system for outputting speech data to the user;

a speech decoding engine for generating an output in response to spoken utterances;

a speech synthesizing engine for generating a synthesized speech output in response to text data;

a dialog manager operatively coupled to the device, the audio input system, the audio output system, the speech decoding engine and the speech synthesizing engine; and

at least one user interface data set operatively coupled to the dialog manager, the user interface data set representing spoken language interface elements and data recognizable by the application of the device;

wherein:

the dialog manager enables connection between the input audio system and the speech decoding engine such that the spoken utterance provided by the user is provided from the input audio system to the speech decoding engine;

the output generated by the speech decoding engine is returned to the dialog

manager;

the dialog manager uses the output generated by the speech decoding engine to search the user interface data set for a corresponding spoken language interface element and data which is returned to the dialog manager when found;

the dialog manager provides the spoken language interface element associated data to the application of the device for processing in accordance therewith;

the application of the device, on processing that element, provides a reference to an interface element to be spoken;

the dialog manager enables connection between the audio output system and the speech synthesizing engine such that the speech synthesizing engine which, accepting data from that element, generates a synthesized output that expresses that element; and

the audio output system audibly presents the synthesized output to the user.

19. An article of manufacture for automatically providing a spoken language interface for a user with respect to at least one external network with which the user interacts, wherein the user possesses a portable spoken language interface device having a data structure for storing one or more user interface data sets used to provide one or more spoken language interfaces, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

requesting a spoken language interface data set from the external network upon discovery of the network;

transferring the spoken language interface data set from the external network to the device; and

loading the spoken language interface data set into the data structure of the device for use by the user interfacing with the external network.